In the Specification

Please replace the second section title on page 1 and paragraphs [0002] and [0003] with the following:

Technical Field ~~of the Invention~~

[0002] This ~~invention~~ disclosure pertains to the field of analysis of sequences of nucleotides and/or amino acids composing living organisms, in particular, analysis of particular mutations of the sequences.

[0003] The ~~invention~~ disclosure also pertains to methods of identification and selection of fragments of sequences of nucleic acids or proteins constituted by and/or comprising motifs having characteristics of specific mutability. The ~~invention~~ disclosure further pertains to pharmaceutical compositions containing the fragments that are useful for treating and/or preventing human, animal and/or plant pathologies or are useful for screening therapeutic compounds.

Please replace paragraph [0014] with the following:

[0014] Moreover, certain web sites such as that of the Los Alamos Library ~~(http://hiv-web.lanl.gov/)~~ provide a large amount of data regarding the alignments of the HIV protein sequences as well as their mutations.

Please replace paragraph [0017] with the following:

[0017] However, these methods do not enable identification of motifs whose mutation possibility is predetermined in relation to the set of sequences analyzed. ~~In the framework of this invention~~, this mutation possibility corresponds to a Boolean state of mutation.

Please replace the first section title on page 5 and paragraphs [0019] through [0021] with the following:

Summary ~~of the Invention~~

[0019] ~~This invention relates to~~ We provide a method for identifying a motif or a combination of motifs having a Boolean state of predetermined mutations in a set of sequences including a) aligning a set of sequences of ordered motifs represented by a single-character code, b) comparing a reference sequence with the set of sequences aligned in step (a), c) identifying motifs not having mutated

simultaneously or motifs having mutated simultaneously at least once on at least one sequence of the set and not having mutated on another sequence of said set.

[0020] ~~This invention~~ We also ~~relates to~~ provide a pharmaceutical composition for treatment of influenza, HIV and hepatitis C including a therapeutically effective amount of the motif or combination of motifs.

[0021] ~~This invention~~ We further ~~relates to~~ provide a method of treating influenza, HIV and hepatitis C including administering a therapeutically effective amount of the pharmaceutical composition.

Please replace paragraphs [0022] and [0023] with the following:

Detailed Description

[0022] ~~This invention provides~~ We provide a new tool to enable finding more durable solutions during therapeutic treatments of pathologies involving pathogenic organisms or human genes having a high degree of mutability.

[0023] ~~The invention~~ We also ~~provides~~ provide for the use of sequences constituted by or comprising the motifs and/or combinations of motifs thereby identified treating or preventing human, animal or plant pathologies, the preparation of therapeutic targets for the screening of said drugs, the docking of a drug on its target, the development of new diagnostic tools in which, for example, the selection of one or more therapeutic agents can be performed as a function of the mutability of the pathogenic organism responsible for the disease of a given patient.

Please replace paragraphs [0028] through [0030] with the following:

[0028] Thus, ~~the invention identifies~~ we identify motifs that did not mutate simultaneously among all of the members of a set of sequences. The identification of such motifs is a major achievement among new pharmacological developments both in terms of therapeutic targets as well as at the level of the searching for new therapeutic compounds, especially in the framework of resistance and multiple-resistances developed by pathogenic organisms which are harmful for both animal species as well as plant species.

[0029] ~~The invention~~ We also ~~pertains to~~ provide for the use of these fragments of sequences constituted by and/or comprising motifs that did not mutate simultaneously for therapeutic targets

that are useful for screening drugs as well as for vaccines directed against pathogenic organisms and, in particular, against pathogenic organisms having a high degree of mutability.

[0030] ~~The invention~~ We further ~~pertains to~~ provide for the use of sequences constituted by and/or comprising motifs that did not mutate simultaneously for compounds useful for preventing and treating human and/or animal pathologies, and in particular pathologies the responsible genes of which have a high degree of mutability.

Please replace paragraphs [0032] through [0034] with the following:

[0032] ~~This invention~~ We, thus, ~~provides~~ provide a new tool for optimizing selection of therapeutic treatments directed against pathogenic organisms with a high degree of mutability or against pathologies due to the appearance of mutations.

[0033] One aspect of the method ~~of the invention~~ for identifying motifs comprises comparing a subset of variants of the same nucleotide or polypeptide sequence of a given pathogenic organism by a reference sequence, for example, a consensus sequence, and then identifying during this comparison the motifs of the sequences which did not mutate simultaneously or the motifs which mutate simultaneously at least once on at least one of the sequences of the subunit and do not mutate on the other sequences of the subunit.

[0034] ~~The invention~~ We more precisely ~~provides~~ provide a method for identifying a motif or a combination of motifs having a Boolean state of predetermined mutation in a set of sequences, comprising:

a) alignment of sequences of ordered motifs represented by their single-character code,

b) comparison of a reference sequence with the set of sequences aligned in step (a),

c) identification of the motifs that did not mutate simultaneously or of the motifs having mutated simultaneously at least once on at least one of the sequences of the set and not having mutated on the other sequences of the set.

Please replace paragraph [0043] with the following:

[0043] Step (b) comprising comparison of sequences of the identification method ~~of the invention~~ advantageously comprises:

– constituting a first numerical matrix A of dimensions NxM in which N designates the number of sequences and M designates the number of motifs of one of the sequences of ~~said~~ the alignment, with the value $A_{i,j}$ being equal to a first value A1 [for example, "0"] when the motif of position i of the sequence j is mutated in relation to the motif of position i of the reference sequence and equal to a second value A2 [for example, "1"] in the other cases,

– constituting two analysis matrices B[[,]] and C of the mutations in which the matrices are:

– a matrix B of unmutated couples, i.e., of couples which did not mutate simultaneously, of dimension MxM, the value $B_{i,k} = B_{k,i}$ being equal:

- to a first value B1 [for example, "0"] when $A_{i,j} = A_k[[.]]_j = A1$ irrespective of the value of j ranging from 0 to N,
- to a second value B2 [for example "1"] in the other cases;

– a matrix C of mutated couples [i.e., of couples that mutate either always, or never simultaneously] of dimension MxM, the value $C_{k,i} = C_{i,k}$ being equal:

- to a second value C1 [for example, "1"] when $A_{i,j} = A_{k,j}$ irrespective of the value of j ranging from 0 to N,
- to a first value C2 [for example, "0"] in the other cases;

– of determining for a set E of positions a coefficient R_E whose value is R_1 [for example, "1"] when the values $B_{i,k}$ are equal to the second value B_2, irrespective of the values of i and k belonging to the set E of the positions, ~~in which i k,~~

– of determining for a set F of positions, a coefficient R_F, the value of which is R_1 [for example, "1"] when the values $C_{i,k}$ are equal to the second value C1, irrespective of the values of i and k belonging to the set F of the ~~position in which i k~~ positions.

Please replace paragraphs [0046] and [0047] with the following:

[0046] The matrix of mutated couples ~~of the invention~~ advantageously makes it possible to identify two motifs having mutated simultaneously at least once on at least one of the sequences of the set and not having mutated on the other sequences of the set.

[0047] ~~The invention~~ We also ~~pertains to~~ provide a way for performing a comparison of the sequences containing the motifs and identifying the motifs thereof, either having mutated

simultaneously at least once on at least one of the sequences of the set and not having mutated on the other sequences of the set and comprising:

– constituting a first numerical matrix A of dimensions NxM in which N designates the number of sequences and M designates the number of motifs of one of the sequences of the alignment, the value $A_{i,j}$ being equal to a first value A_1 [for example, "0'] when the motif of position i of the sequence j is mutated in relation to the motif of position i of the reference sequence and equal to a second value A_2 [for example, "1"] in the other cases,

– constituting two analysis matrices B[[,]] and C of the mutations M in which this matrix is:

– a matrix B of unmutated couples, i.e., couples which did not mutate simultaneously, of dimension MxM, the value $B_{i,k} = B_{k,i}$ being equal:

- to a first value B1 [for example, "0"] when $A_{i,j} = A_{k,j} = 0$ irrespective of the value of j ranging from 0 to N,
- to a second value B2 [for example, "1"] in the other cases;

– a matrix C of mutated couples [i.e., couples that mutate either once simultaneously or never] of dimension MxM, the value $C_{i,k} = C_{k,i}$ being equal:

- to a second value C1 [for example, "1"] when $A_{i,j} = A_{k,j}$ irrespective of the value of j ranging from 0 to N,
- to a first value C2 [for example, "0"] in the other cases;

– of determining for a set E of positions a coefficient R_E, the value of which is R1 [for example, "1"] when all of the values $B_{i,k}$ are equal to the second value B2, irrespective of the values of i and k belonging to the set E of said positions, ~~in which i j,~~

– of determining for a set F of positions a coefficient R_F the value of which is R1 [for example, "1"] when all of the values $C_{i,k}$ are equal to the second value C2, irrespective of the values of i and k belonging to the set F of said ~~positions, in which i j~~ positions.

Please replace paragraph [0052] with the following:

[0052] ~~The invention pertains to identifying~~ We provide motifs belonging to pathogenic agents, the nucleic acid and/or polypeptide sequences of which are capable of having mutations.

Please replace paragraphs [0054] and [0055] with the following:

[0054] Thus, according to a particular aspect ~~of the invention~~, the subset of extracted sequences comprises the polypeptide sequences of the different variants of the neuraminidase of the flu virus.

[0055] According to another particular aspect ~~of the invention~~, the subset of extracted sequences comprises all of the polypeptide sequences of the different variants of the hemagglutinin of the flu virus.

Please replace paragraph [0057] with the following:

[0057] The method for identifying motifs ~~of the invention~~ is not limited solely to the domain of pathogenic agents. Sets of sequences having motifs which did not mutate simultaneously, or in contrast had mutated together at least once on at least one of the sequences of the set and had never mutated on the other sequences of the set are also presented in other pathologies such as, for example, pathologies in the field of cancer research.

Please replace paragraph [0060] with the following:

[0060] ~~The invention~~ We also ~~includes identifying~~ identify motifs described above for selecting fragments of sequences constituted by and/or comprising motifs that did not mutate simultaneously for vaccines.

Please replace paragraph [0063] with the following:

[0063] The application of the method ~~of the invention~~ to subsets of variant sequences of the protein sequences of pathogenic sequence makes it possible to trap these mutant virus:

– either it mutates but, in this case, it is no longer functional;

– or it does not mutate, but then the antibodies produced by the vaccine will be capable of destroying it.

Please replace paragraphs [0067] through [0069] with the following:

[0067] ~~The invention also pertains to the~~ We use ~~of~~ fragments of sequences constituted by and/or comprising nucleotide and/or peptide motifs of the analyzed sequences that did not mutate simultaneously for a vaccine.

[0068] ~~The invention~~ We also ~~includes~~ use a method for identifying motifs or combination of motifs that did not mutate simultaneously to develop diagnostic tools. ~~The invention~~ We further ~~includes~~ include use of such an identification method to fragments of sequences constituted by and/or comprising motifs having mutated simultaneously for diagnostic tests.

[0069] The method ~~of the invention~~ also makes it possible to construct a database which constitutes a decision-making tool, for example, for determining by the physician of the administration of antiviral therapies to a given patient.

Please replace paragraphs [0071] through [0075] with the following:

[0071] The drug-mutated amino acid relationship demonstrated in this manner is very useful for improving treatment. For example, with regard to HIV, comparison of the peptide motifs is performed on three subsets of a protein database, pertaining to reverse transcriptase, protease and integrase (~~http://hiv-web.lanl.gov/~~).

[0072] The comparison of the sequences belonging to the subsets comprising from about 300 to about 8000 sequences or fragments of the sequences of each of these three proteins enables application of the method ~~of the invention~~ to identify combinations of amino acids that did not mutate simultaneously.

[0073] Thus, the method ~~of the invention~~ makes it possible to identify the mutations induced under the pressure of selection.

[0074] The aspect ~~of the invention~~ comprising comparison with the drug resistances enables selection of a combination of drugs such that the amino acid mutations capable of being induced by each of the antiviral agents, capable of conferring resistance on the various drugs involved in this combination (fewer than ten), are not produced simultaneously. Identification of such motifs enables selection of a drug combination which disfavors the appearance of more than one mutation at a time, thereby closing the door to multiple resistances. The practitioner can then use the information obtained by applying this method, for example, to isolated viral sequences or viral sequences deduced from the isolated viral genome, of a given patient to ensure that the envisaged multi-drug therapy is in fact the most effective possible. With the identification of a first mutation excluding the two others, a selected three-agent therapy thereby enables the two remaining antiretroviral agents to continue to be effective.

[0075] The aspect of identification of peptide regions not having mutated simultaneously also provides valuable assistance in the case of the appearance of resistances in already treated patients. The method ~~according to the invention~~ can, for example, be applied to the subsets of polypeptide sequences among which is included that or those deduced from the sequencing of the isolated viral genome of the patient. Thus, if this genotyping reveals a mutation responsible for resistance, the method of identification of peptide motifs not having mutated allows implementation of a multiple-therapy regimen designed to maintain the selection pressure on the mutation. The molecule identified in this manner can be accompanied by two or three antiretroviral agents which target domains of the protein not capable of mutating at the same time as the zone that mutated.

Please replace paragraphs [0077] through [0085]with the following:

[0077] ~~The invention~~ We also ~~provides~~ provide for identification of a set of genes or a set of noncoding sequences of motifs not having mutated simultaneously. Identification of such motifs enables selection of genetic regions that can have physical or functional interactions on the overall genome.

[0078] Another aspect ~~of the invention~~ relates to a method for identifying motifs and combinations of motifs for selecting fragments constituted by and/or comprising motifs not having mutated simultaneously for the preparation of therapeutic targets.

[0079] Still another aspect ~~of the invention~~ pertains to the use of fragments of sequences constituted by and/or comprising motifs either having mutated at least once on at least one sequence of the set and not having mutated on the other sequences of the set for the preparation of therapeutic targets.

[0080] ~~The invention~~ We also ~~relates the~~ use ~~of~~ motifs or combinations of motifs identified in this manner for preparing therapeutic targets that are useful for screening new therapeutic compounds to prevent and/or treat human, animal or plant pathologies. Thus, the preparation, after having identified motifs not having mutated simultaneously, or sequence fragments containing them, enables preparation of a therapeutic target against which will be tested therapeutic compounds directed against the pathogenic organism and especially therapeutic compounds against which the wild pathogenic organism can not develop resistance mutations.

[0081] The selection of fragments constituted by and/or comprising motifs not having mutated simultaneously is, thus, useful for the preparation of diagnostic tools since it is not always easy to

detect rapidly a certain type of or subtype of pathogenic organism, because the identification of peptide motifs ~~according to aspects of the invention~~ enables preparation of fragments of peptides comprising the motifs most representative of a subtype of a pathogenic organism. These fragments are then used in detection tests such as, for example, immunoenzyme tests.

[0082] This application ~~of the invention~~ comprises identifying a set of motifs indispensable for the function of a protein of a human, animal or plant organism or of a pathogenic organism. These motifs can constitute, for example, a subset of amino acids known to play an important role in the function of the targeted protein. The motifs identified in this manner are advantageously contiguous motifs of the genetic sequence and represent a linear sequence of the gene. The motifs identified are advantageously motifs noncontiguous on the linear sequence of the gene. They can then be useful for completing three-dimensional analysis studies to confirm a possible nonlinear spatial proximity of the motifs. The method ~~of the invention~~ can then include a new supplementary step (g) after the step (e) of identification of the motifs, the step comprising comparing the motifs with the three-dimensional structural data of these proteins such as the amino acids involved in the catalytic site and/or in the sites linked by noncompetitive inhibitors. This latter comparison produces a list of amino acids involved in the protein function and not having mutating together.

[0083] ~~The invention~~ We also ~~uses~~ use fragments of sequences constituted by and/or comprising peptide motifs having mutated simultaneously for the development of diagnostic tools. The method for the identification of peptide regions defines the most representative peptides of a subtype. Once they are identified, these peptides are used in detection tests known in the art, such as, for example, immunoenzyme tests of the ELISA type.

[0084] The search for peptides representing a subtype of a particular type is performed as indicated above. It is a question of finding peptide antigens capable of being recognized by a particular serum containing or not containing the antibodies of a particular subtype. The method ~~according to the invention~~ can be applied to any databank of sequences. The results are compared by subtypes and the theoretical peptide combination the most representative of a particular pathogenic type is thereby identified. The peptides identified in this manner are synthesized and tested immunologically against a collection of serums.

[0085] ~~The invention exhibits its~~ Our methods exhibit their value especially when it is used for the identification either of motifs having mutated once together or not having mutated, from a large

number of sequences comprising a large number of motifs to select the sequences of motifs useful for the various applications envisaged above.

Please replace paragraph [0086] with the following:

[0086] To illustrate the method for the identification of motifs ~~of the invention~~, the example below shows the different matrices constituted in a comparison of motifs performed on a subset of eight sequences based on the reference sequence S V R L G H K D E V (SEQ ID NO: 1). The peptides that follow are shown in SEQ ID NOS 1-9, respectively, in order of appearance.

POSITIONS	0 1 2 3 4 5 6 7 8 9
Reference sequence (consensus)	S V R L G H K D E V

Subset of sequences	Alignment
SEQ ID NO. 2	S R R L G H K D E V
SEQ ID NO. 3	S V R L G H K L E V
SEQ ID NO. 4	S R D L G H K D E V
SEQ ID NO. 5	S V R L G H L D V V
SEQ ID NO. 6	S V D L G H K T E V
SEQ ID NO. 7	S K R L G H K D E V
SEQ ID NO. 8	S V R L G H G D G V
SEQ ID NO. 9	S V R L G H K S E V